「정보시스템연구」 제31권 제2호
한국정보시스템학회
2022년 6월, pp. 239～261
http://dx.doi.org/10.5859/KAIS.2022.31.2.239

Understanding the Changes in Tourists' Opinions in the Era of the COVID-19*

Chernyaeva, Olga** · Ziyan, Yao*** · Hong, Taeho****

〈Content〉

Ⅰ. Introduction

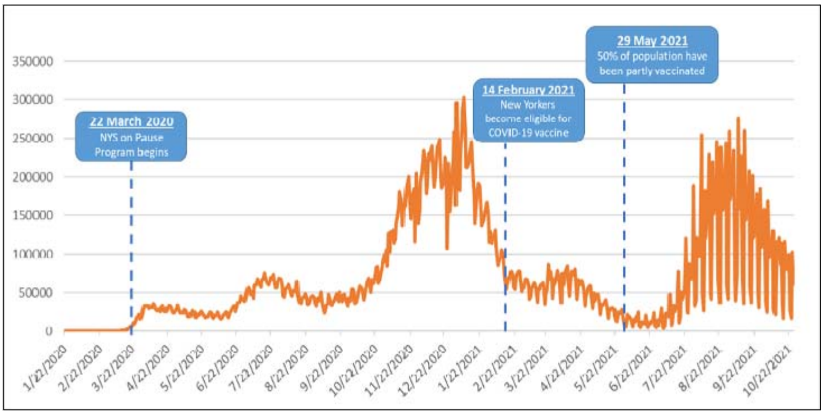
COVID-19 is a newly discovered coronavirus that has spread worldwide and was declared a pandemic after WHO's Pandemic Declaration on March 11, 2020. To prevent transmitting a global pandemic across national and continental boundaries, governments have enacted laws and regulations restricting people's traveling between countries and within the country (Jamal and Budke, 2020). As a result, the notions of "lockdown" and "social distancing" have entered our common lexicon (Long, 2020). The COVID-19 pandemic profoundly reshaped the global tourism landscape and significantly decreased tourist arrivals, especially in the case of mass tourism attractions with large volumes of tourists before the pandemic decreased. According to the UNWTO (United Nations World Tourism Organization) statement, pandemic-related restrictions have reduced the national tourist market to a 44% decrease from 2019, meaning 180 million tourists (UNWTO, 2020).

* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019S1A3A2098438)
** Business School, Pusan National University, misslelka@pusan.ac.kr (Lead author)
*** Business School, Pusan National University, yaoziyan@pusan.ac.kr
**** Business School, Pusan National University, hongth@pusan.ac.kr (Corresponding author)

Furthermore, as the spread of COVID-19 and travel restrictions from each government has continued, all industries related to tourism, including hotels, restaurants, and shopping malls, are “already facing collapse” or “in a fight for survival” (Jiang and Wen, 2020). Also, the pandemic has led to changes in tourists' habits and behaviors.

Since the outbreak of COVID-19, researchers have generally observed the adverse effects of the pandemic on tourism. Chua et al. (2020) state that after the COVID-19 outbreak, tourists started to avoid popular tourist destinations and prefer non-crowded attractions. Moreover, after the COVID-19 outbreak, adverse responses to crowding become more common (Park et al., 2021, Yoon et al., 2021). However, previous studies have shown that even before COVID-19, crowding negatively impacted tourists' opinions, satisfaction, and experience (Sharp et al., 2015, Li et al., 2017).

After the mass vaccination, when the COVID-19 restrictions began to weaken, more and more people visited the tourist attractions. Therefore, the restarting of international travel is affected by the COVID-19 recovery strategy and mass vaccination (Bernal et al., 2021). Furthermore, since vaccination can reduce tourists' anxieties and fears caused by the COVID-19 pandemic, Gursoy et al. (2021) stated that increasing vaccination rates may significantly impact hospitality recovery. Therefore, recent researchers mainly concentrate on issues such as vaccine tourism (Williams et al., 2021) and determining factors that influence pre-travel vaccination intention (Suess et al., 2022; Pavli and Maltezou, 2022). However, there is a lack of studies from the post-travel perspective on the influence of mass vaccination and COVID-19 outbreak on tourists' opinions about attractions. Therefore, the novelty of our study is that we provide a comparative analysis of tourists' opinions from



<Figure 1> The timeline of new cases and key events related to COVID-19 in the US

before the COVID-19 outbreak to when more than half of the local population had been vaccinated. This study aims to explore and compare changes in tourist opinion caused by the COVID-19 pandemic and the vaccination of more than half of the population from the post-travel perspective based on online reviews.

Moreover, there are differences in the number of COVID-19 confirmed patients by country, which may result in different effects in each country simultaneously (Fanelli and Piazza, 2020). Therefore, we consider it essential to understand the evolution of tourists' opinions and mindsets in a particular country during the pandemic to adapt travel operations to tourists' necessities. Thus, the present study aims to understand and reveal changes in the US tourists' opinions before/after the COVID-19 outbreak and after the vaccination of more than half of the country's population. In the local context of the US, New York, there have been a few critical events related to COVID-19. The timeline of new cases and critical events related to COVID-19 in the US are shown in Figure 1. In the US, the first case of COVID-19 was confirmed on January 21, 2020, and on March 22, 2020, New York City on Pause program began (www. nytimes.com). In addition, on February 14, 2021, New Yorkers became eligible for the COVID-19 vaccine, and on May 29, 2021, 50% of the US population was partly vaccinated (www.covid19.who.int).

According to the study of Zaman et al. (2021), the desire of tourists to travel depends on the number of vaccinated people, and it is also reinforced by the study of Ram (2022) that the more vaccinated people, the more people feel safe. Since the moment of vaccinating half of the local population is a transitional moment from the minority of the vaccinated to the majority of the vaccinated, in our study, we considered the day when more than half of the population was vaccinated as a key date. The data was collected from TripAdvisor.com regarding the tourists' concerns in the form of online reviews and the popularity of tourist attractions in the US (New York). To analyze tourists' opinions of tourist attractions, we perform sentiment analysis, emotion analysis, and topic modeling through Latent Dirichlet Allocation (LDA) by summarizing topics into five dimensions: management, scenery, price, suggestion, and safety, then, based on the topic modeling results. Therefore, in this study, we will answer the following research questions:

1) How have the tourists' sentiments about attractions changed from before the COVID-19 outbreak to when more than half of the local population had been vaccinated?

2) How have the tourists' emotions about attractions changed from before the COVID-19 outbreak to when more than half of the local population had been vaccinated?

3) How have the topic dimensions of the

tourists' attractions reviews changed from before the COVID-19 outbreak to when more than half of the local population had been vaccinated?

The theoretical contribution of our study is that we introduced a methodology to evaluate tourists' opinion changes before/after the COVID-19 outbreak and the after the vaccination of more than half of the population by text mining techniques and showed the differences between each period's opinion is statistically significant level. Moreover, our research may help the industry better to understand the changes in tourists' opinions about attractions and improve tourist services for higher customer satisfaction.

The structure of this paper is as follows. First, Section 2, Literature Review, the role of online reviews in tourism, describes the impact of COVID-19 on tourism and methods to evaluate tourist opinions. Then, Section 3, Research Frameworks and Experiments, presents our research framework and describes the methodology and techniques. Next, Section 4 provides the results and discussions of our study. Finally, Section 5 discusses our findings, contributions, and limitations.

Ⅱ. Literature Review

2.1 The Role of Online Reviews in Tourism

The advancement of information and communication, which has emerged from the spread of various electronic devices and the Internet worldwide, improves the quality of life since the industry has widely adopted online business. The key to this change is the ease of access to the Internet due to the spread of mobile phones and the convenience of sharing opinions through online reviews. Users can freely and easily access information and exchange opinions about products or services purchase experience on an unprecedented scale in real-time. Nowadays, online reviewers can share their experiences and emotions through the reviews written on online platforms, influencing potential consumers' exploration of information and purchase (야오즈옌 등, 2020). Also, before a trip, tourists search for reliable and accurate information about the destination to reduce the risks of failure. Therefore online travel reviews play a significant role in potential tourists' decision-making (Mauri and Minazzi, 2013; Jacobsen and Munar, 2012). One of the most popular travel sites is Tripadvisor. Tripadvisor is the world's largest travel platform that helps travelers plan their trips or share their experiences with others (Tripadvisor, 2022). Tripadvisor's travel guidance shows where to go, what to do, and where to stay based on opinions from tourists who have been there before. The advantage of TripAdvisor is that users can easily access, search and share their opinions and travel experiences by uploading images and posting reviews. Therefore the role of reviews is critical for the TripAdvisor platform (Saydam et al., 2022). In 2020 tourists have written more than 830 million reviews on the TripAdvisor platform (Tripadvisor, 2020), and 83% of users think it is essential to read and reference reviews before a trip (Feinstein, 2018). In addition, the spread of the e-commerce market has been accelerated by the epidemic of COVID-19 (Al-maaitah et al., 2021). Utilizing online reviews in e-commerce helps customers make better decisions such as search and purchase processes(Cheng and Ho, 2015; 이연수 등, 2021). Therefore, understanding and utilizing online reviews is vital as this trend becomes a new part of economic growth (Li et al., 2013).

The COVID-19 pandemic has brought global-scale health crises, and tourism industries have become one of the victims of this crisis (Kim et al., 2020). Prospective tourists' information search changes according to their risk perception of the COVID-19 pandemic (Im et al., 2021). Since tourists search for critical information about attractions' regulations and safety levels through reviews, especially in COVID-19, interpreting online reviews and using them in decision-making has been more critical than ever (Gretzel et al., 2020). Therefore, our study focused on the top most popular tourist attractions to understand how tourists' attitudes have changed with COVID-19 and what tourists have become concerned about when evaluating an attraction.

2.2 COVID-19 and Tourism

The COVID-19 pandemic caused global-scale health and economic and social crises (Yoon et al., 2021). Governments have enacted laws and regulations to prevent transmitting a global pandemic across national and continental boundaries (Jamal and Budke, 2020). For instance, the regulations such as border closures, travel bans, and quarantine directives, significantly decreased tourist arrivals and reshaped the global tourism landscape (Bernal et al., 2021; Suess et al., 2022). Hospitality and tourism industries suffered huge losses from the COVID-19 pandemic; numerous tourism-related enterprises scaled back or shut operations (Williams et al., 2021). For example, in the US, compared with 2019, in March 2020, sales of tourism industries, including travel agencies, airlines, cruises, etc., decreased by 85% (Im et al., 2021). Moreover, pandemic-related restrictions have reduced the national tourist market to a 44% decrease from 2019, meaning 180 million tourists (UNWTO, 2020).

However, along with the closure of borders and difficulties with international transfers, the COVID-19 outbreak positively impacted local tourism development. People started to search for domestic trips such as short-term trips to nature, trips to new places inside of the city or state, or tours to popular attractions that used to be too crowded before the COVID-19 outbreak (Mary and Mahmoud, 2022). For example, according to the study by Im et al. (2021), in South Korea, the number of confirmed COVID-19 cases and news about the pandemic impact online information search for tourist attractions.

As an effective recovery strategy against the COVID-19 crisis, mass vaccination could help and significantly impact the tourism industry's recovery. The restarting of international travel is affected by the COVID-19 recovery strategy and mass vaccination (Bernal et al., 2021). Therefore, after the mass vaccination, the anti-covid restrictions were less severe, borders opened, and local and international tourists started to visit the tourist attractions (Suess et al., 2022). Furthermore, since vaccination can reduce tourists' anxieties and fears caused by the COVID-19 pandemics, increasing vaccination rates may significantly impact hospitality recovery (Gursoy et al., 2021). Recently, in the hospitality research area, the topics such as vaccine tourism (Williams et al., 2021) and pre-travel vaccination intention (Suess et al., 2022; Pavli and Maltezou, 2022) have become hot. However, there is a lack of studies from the post-travel perspective on the influence of mass vaccination and COVID-19 outbreak on tourists' opinions about attractions.

2.3 Methods for Measuring Tourist Opinions

Researchers use two methods for evaluating tourists' opinions: surveys and tourists generated content such as reviews or tweets. The most common techniques to evaluate tourist opinion from online reviews using text mining techniques are sentiment analysis, emotion analysis, and topic modeling (Alamoodi et al., 2020; Ozturk and Ayvaz, 2018; Xie et al., 2021). The description of each of these techniques is as follows.

2.3.1 Sentiment and Emotion Analysis

Sentiment analysis is used to identify and extract the positive and negative parts of the text (Hu et al., 2012). Since various electronic devices and platforms have been developed, online reviews have sharply increased (야오즈옌 등, 2020). Therefore, sentiment analysis of online reviews has become a popular research area (Kumar and Teeja., 2012). There are several methodologies to calculate sentiment scores, such as lexicon-based, machine learning-based, and hybrid methods (Alamoodi et al., 2020). Lexicon-based methods classify textual content using a dictionary of manually labeled lists of words with a sentiment score (Hu and Liu, 2004; Baccianella et al., 2010; Ozturk and Ayvaz, 2018). Machine learning methods utilize machine learning algorithms to classify textual content (Pak and Paroubek,

<Table 1> Summary of Sentiment and Emotion Analysis Research

Analysis	Approach	Previous studies
Sentiment	Lexicon-based	Hu and Liu (2004), Baccianella et al. (2010), Ozturk and Ayvaz (2018)
	Machine learning-based	Pak and Paroubek (2010), Shoeb and Ahmed (2017)
	Hybrid	Govindarajan (2013), Mukwazvure and Supreethi (2015)
Emotion	Lexicon-based	Poria et al. (2013), Mohammad and Turney (2013), Srinivasan et al. (2019)
	Machine learning-based	Hasan et al. (2014)
	Hybrid	Yang et al. (2012), Khanpour and Caragea (2018)

2010; Shoeb and Ahmed, 2017). Finally, hybrid methods combine lexicon and machine learning approaches (Govindarajan, 2013; Mukwazvure and Supreethi, 2015). In previous studies, the lexicon-based method was used for calculating sentiment scores because this method is easy to apply and widely used in other research (Ma et al., 2018). Also, Shafqat and Byun (2020) analyzed tourist reviews through a lexicon-based method and calculated sentiment scores, then developed a system that recommends tourist destinations with high ratings and positive sentiment scores.

Previous researchers described emotion analysis as a part of the larger field of sentiment analysis (Hakak et al., 2017). For example, in Oxford Dictionary, "emotion" was defined as: "A strong feeling deriving from one's circumstances, mood, or relationships with others." Other researchers described emotion analysis as finding subjective feelings and thoughts (Das and Bandyopadhyay, 2014). Therefore we can summarize that emotion analysis is analyzing text data and identifying expressed emotions. However, emotions are different from sentiment analysis. Sentiments represent a polarity of emotion: positive, negative, or neutral, but emotions represent a reaction that people have to a particular event, such as happiness, sadness, anger, etc. (Munezero et al., 2014). For instance, if a tourist writes "I am happy," then the sentiment is "positive," and the emotion is "happy" (Sailunaz and Alhajj, 2019). A summary of research that applied sentiment and emotion analysis is shown in Table 1. For emotion analysis, same with sentiment analysis, previous researchers used three methods: lexicon-based, machine learning-based, and hybrid (Medford et al., 2020). In this study, we adopted the lexicon-based method due to this method was widely used by previous researchers.

2.3.2 Topic Modeling

Topic modeling is an unsupervised learning technique that learns variables according to the frequency of simultaneous occurrences between words. LDA (Latent Dirichlet Allocation) is a widely used method for analyzing text data. Using LDA, review topics can be extracted and classified by related topics (Blei et al., 2003; 홍태호 등, 2018). Several studies have analyzed online reviews on online e-commerce platforms by applying text mining techniques. Carracedo et al. (2020) used the clustering technique to capture changes in the circumstance during COVID-19. He analyzed the changes in the business environment under the circumstance of COVID-19 and drew three different areas: "public health," "consumption habits," and "economic effects." Xie et al. (2021) applied LDA (Latent Dirichlet Allocation) to understand the public response to COVID-19 on TripAdvisor.com.

In this study, topics and corresponding keywords are extracted by LDA, and various topics discussed concerning COVID-19 on TripAdvisor.com are discovered. Tong and Zhang (2016) used LDA to extract topics from Wikipedia and Twitter even with a small number of datasets. Other researchers applied LDA and combined it with the dimension of a specific field. For example, Poria et al. (2016) applied LDA-based topic modeling techniques to extract three perspectives of the hotel industry ('location,' 'service,' 'value'). Meanwhile, before applying LDA, finding an optimal number of topics is crucial for the performance of LDA. Some researchers used perplexity to determine the optimal number of topics (Zhao et al., 2015; Hong et al., 2019). Hong et al. (2019) found an optimal number of topics using perplexity with 5-fold cross-validation when analyzing the hotel reviews.

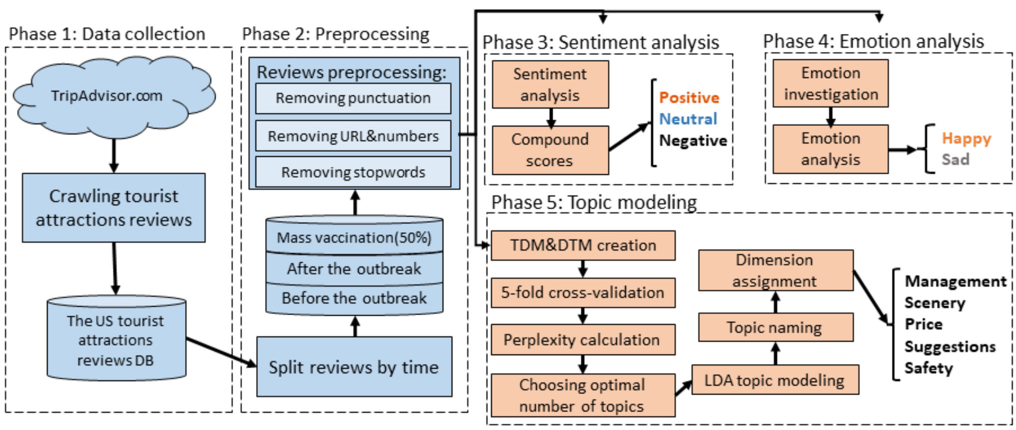
Ⅲ. Research Framework and Experiment

As shown in Figure 2, our research framework consists of five phases: "Phase 1: Data collection", "Phase 2: Preprocessing", "Phase 3: Sentiment analysis", and "Phase 4: Emotion analysis," and "Phase 5: Topic modeling". For example, in Phase 1: Data collection, we collected review data from TripAdvisor.com. We have chosen the top tourist attractions of New York, United States, for target data. TripAdvisor.com provided the list of top-10 tourist attractions: Brooklyn Bridge, Central Park, Empire State Building, Manhattan Skyline, One World Observatory, Statue of Liberty, The High Line, The Metropolitan Museum of Art, The National 9/11 Memorial Museum, Top of the Rock. We collected tourists' reviews by web scraping in Python. In Python, we used the Beautiful Soup library to automatically extract large amounts of review data from the TripAdvisor.com website. The collected reviews are from

February 1, 2019, to October 31, 2021. As a result, we collected 24,013 New York tourist attractions reviews in November 2021 as this research dataset. In our research, we only process reviews written in English.

In Phase 2: Preprocessing, we split data by period based on key events related to COVID-19 in the US (New York), as shown in Figure 1. The first dataset period is before the COVID-19 outbreak from February 1, 2019, to March 22, 2020, and the second dataset is reviews created after the COVID-19 outbreak from March 22, 2020, to May 29, 2021. Furthermore, the third dataset is reviews created after May 29, 2021, when 50% of the US population had been partly vaccinated (www.covid19.who.int). Partly vaccinated people are people who got only one shot of the COVID-19 vaccine. Fully vaccinated people are people who complete vaccination for the COVID-19. After constructing three attractions' review databases, preprocessing procedures are applied to review datasets. Preprocessing procedures are as follows. First, we divided reviews into single words. Second, we removed stopwords such as website links, numbers, symbols, and punctuation because stopwords hinder the reliability of experiment results and make it hard to interpret the results (Aggarwal and Zhai, 2012).

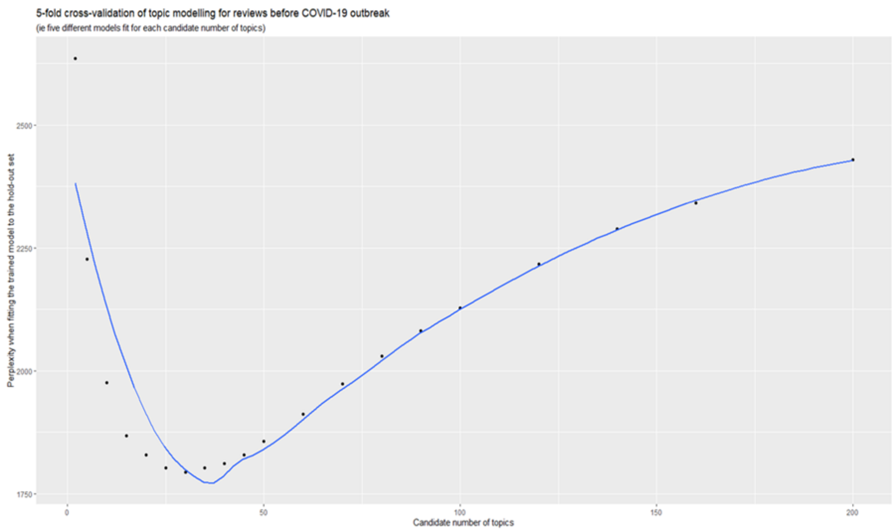
We provided sentiment and emotion analysis through Python using the lexicon-based method. In phase 3, for sentiment analysis, we used the SentiWordNet package. SentiWordNet is a resource containing an opinion lexicon extracted from the WordNet database (Ohana and Tierney, 2009). Each term is associated with numerical scores indicating positive, neutral, and negative sentiment scores in the WordNet database. SentiWordNet was made publicly available for research purposes of text content in English. In phase 4, for emotion



<Figure 2> Research framework to explore and compare changes in tourist opinion during the COVID-19

analysis, we used the text2emotion package. Text2emotion is the python package developed by (Diaz et al. 2018) to recognize the emotions embedded in the text data, such as Happy, Angry, Sad, Surprise, and Fear. The goal of tourist attraction organizations is to satisfy tourists, to leave them with positive emotions after visiting so that the tourist would later return to the attraction or recommend it to friends or social networks (Bigné et al., 2005). Tourists or customers experiencing more positive sentiments and emotions showed an increased level of satisfaction and more favorable behavioral intentions, meaning loyalty and willingness to pay more (Egger et al., 2019). The positive emotion is happiness, and the polar emotion of happiness is sadness (Lynott and Coventry, 2013). Moreover, many researchers consider happiness and sadness are also considered to be basic emotions, as well as being two of the most commonly used terms related to affect and satisfaction (Lynott and Coventry, 2013; Batty and Taylor, 2003; Ekman, 1992). Therefore, in our study, to assess the opinion of tourists about attractions, we focused our attention on two polar emotions: happiness and sadness.

In the last phase, Phase 5, we applied LDA on review datasets. We calculated the perplexity by the number of topics combined with a 5-fold cross-validation for each period dataset to find an optimal number of topics. We calculated the perplexity as the number of topics increased. The number of topics can be selected where the slope of the graph becomes smooth. The example of perplexity with a 5-fold cross-validation plot for reviews before the outbreak is shown in Figure 3.



<Figure 3> An Example of Perplexity with 5-fold Cross-validation

A number of topics with the lowest k perplexity value are optimal. For example, the results of the optimal number of topics for the before outbreak reviews dataset is 35, for the after outbreak reviews dataset is 20, and for the mass vaccination reviews, the dataset is 20 also. Keywords for each topic were extracted by applying LDA to each dataset. After, we named each topic and assigned a dimension. According to the study by Ren and Hong (2017), we summarized the topics into four dimensions (Management, Scenery, Price, and Suggestion). Since the COVID-19 pandemic changed our lives and people started to think more about health care, we also highlighted the fifth dimension - Safety (Ridhwan and Hargreaves, 2021).

Ⅳ. Results

4.1 Sentiment and emotion analysis results

The results of sentiment and emotion analyses are shown in Table 2 and Table 3. To answer the first research question: “How have the tourists' sentiments about attractions changed from before the COVID-19 outbreak to when more than half of the local population have been vaccinated?” we calculated mean sentiment analysis scores and did an ANOVA test. The mean sentiment scores of the time before the COVID-19 outbreak are 0.844, the mean sentiment scores of the time after the COVID-19 outbreak till mass vaccination are 0.918, and the mean sentiment scores of the time after vaccination, more than half of the local population are 0.822. Based on the ANOVA test results shown in Table 2, we can see the statistically significant differences between the means of tourists' sentiments before, after the COVID-19 outbreak, and after the vaccination of more than half the local population. To understand the different results among the periods in detail, we also provide an ANOVA post hoc test. In case when an analysis of variance (ANOVA) is significant, a post hoc test is used to detect specific differences between three or more group means (Brown, 2005). From Table 3, we can see post hoc test results showing us the statistically significant differences between the means of tourists' sentiments before and after the COVID-19 outbreak and tourists' sentiments after the COVID-19 outbreak and after the vaccination of more than half the local population. However, differences between the means of tourists' sentiments before the COVID-19 outbreak and after the vaccination of more than half the local population are not significant. Thus, our study proved that the sentiments of tourists about attractions changed from the time before the COVID-19 outbreak to when more than half of the local population was vaccinated. Also, in our study, we did not find a significant difference in the sentiments of tourists before the pandemic and after the vaccination of half of the local population.

The answer to the second research question: “How have the tourists' emotions about attractions changed from before the COVID-19 outbreak to when more than half of the local population have been vaccinated?” is shown in Table 2. The mean happy emotion scores of the time before the COVID-19 outbreak are 0.247, the mean happy emotion scores of the time after the COVID-19 outbreak till mass vaccination are 0.255, and the mean happy emotion scores of the time after vaccination, more than half of the local population are 0.256. In the case of sad emotions, the mean sad emotion scores of the time before the COVID-19 outbreak are 0.163, the mean sad emotion scores of the time after the COVID-19 outbreak till mass vaccination are 0.152, and the mean sad emotion scores of the time after vaccination more than half of the local population are 0.155. After the COVID-19 outbreak, happiness tends to increase over time. However, in the case of sadness, after the outbreak, sadness tends to decrease. Based on the ANOVA test results shown in Table 2, we can see the statistically significant differences between the means of tourists' happy and sad emotions before, after the COVID-19 outbreak, and after the vaccination of more than half the local population.

Moreover, from Table 3, we can see the results of the post hoc test that show us the statistically significant differences between the means of tourists' happy and sad emotions before and after the COVID-19 outbreak and tourists' happy and sad emotions after the COVID-19 outbreak and after vaccination more than half the local population. However, differences between the means of tourists' happy and sad emotions before the COVID-19 outbreak and after the vaccination of more than half the local population are not significant. As a result, we proved that tourists' happy and sad emotions about attractions changed from before the COVID-19 outbreak to when more than half of the local population was vaccinated. Also, same with sentiments in our study, we did not find a significant difference in tourists' happy and sad emotions before the pandemic and after the vaccination of half of the local population.

<Table 2> ANOVA results of sentiment and emotion analysis

		Mean scores			ANOVA
		Before	After	Mass vac.	p-value
Sentiments		0.844	0.918	0.822	0.00*
Emotions	Happy	0.247	0.255	0.256	0.00*
	Sad	0.163	0.152	0.155	0.00*

Significance levels: * $p < 0.01$

<Table 3> ANOVA post-hoc test results of sentiment and emotion analysis

	p-value
Sentiments	
Before vs. After	0.001*
Before vs. Mas vac.	0.89
After vs. Mas vac.	0.001*
Happy Emotions	
Before vs. After	0.001*
Before vs. Mas vac.	0.852
After vs. Mas vac.	0.001*
Sad Emotions	
Before vs. After	0.001*
Before vs. Mas vac.	0.409
After vs. Mas vac.	0.073**

Significance levels: * $p < 0.01$; ** $p < 0.1$

<Table 4> An example of topic naming and dimension assignment for the period before COVID-19

Topic ID	Dimension	Topic	Top 10 words
1	Scenery	Architecture	Architecture, art, experience, cultural, keep, picture, history, enter, building, exhibition
2	Safety	Cleanliness and refund	Place, tour, clean, detail, line, crowded, email, refund, ticket, time
…	…	…	…
29	Price	Booking ticket	Ticket, book, people, expensive, dollar, tour, museum, price, sale, euro
30	Management	Service	Headphone, architecture, audio, inside, time, people, guide, carrier, thought, outside

4.2 Topic modeling results

In our study, to answer the third question: “How have the topic dimensions of the tourists' attractions reviews changed from the time before the COVID-19 outbreak to when more than half of the local population have been vaccinated?”, we named each topic for each period and assign one of five dimensions: Management, Scenery, Price, Suggestion, and Safety. The example of topic naming and dimension assignment for the period before the COVID-19 outbreak is shown in Table 4.

After, we count the number of each dimension for each period. The result is shown in Table 5. As we can see from the result, the dimension “Safety” was not present during the period before the outbreak. However, in the period after the outbreak and during the mass vaccination dimension, “Safety” appears. These results mean that tourists started to care about safety due to the COVID-19 pandemic (Ridhwan and Hargreaves, 2021). Thus, our study proved that topics of tourists' opinions

<Table 5> The results of topic modeling

	Before outbreak	After outbreak	Mass vaccination
Optimal topics number	35	20	20
Management	8	4	5
Scenery	13	3	3
Price	10	8	7
Suggestions	3	3	4
Safety	0	2	1

about attractions also changed from before the COVID-19 outbreak to when more than half of the local population was vaccinated.

Since the COVID-19 pandemic caused global-scale health and economic and social crises (Yoon et al., 2021), as the spread of COVID-19 and travel restrictions from each government have continued, all industries related to tourism faced collapse (Jiang and Wen, 2020). Furthermore, previous researchers stated that after the COVID-19 outbreak, adverse responses to crowding became more common (Park et al., 2021; Yoon et al., 2021). Also, Gursoy et al. (2021) stated that vaccination could reduce tourists' anxieties and fears caused by the COVID-19 pandemics, and increasing vaccination rates may significantly impact hospitality recovery. Our study complemented past studies, proving that tourists' opinions about attractions changed from before the COVID 19 outbreak to when more than half of the local population was vaccinated. Furthermore, we proved that positive sentiments and emotions increased after the COVID-19 outbreak and decreased by more than half of the local population after vaccination.

Moreover, the dimension 'Safety' started to appear in tourists' opinions about attractions reviews only in the period after the outbreak and during the mass vaccination. These results mean that tourists started to care more about safety due to the impact of the COVID-19 pandemic. Also, based on analyzing tweets, Ridhwan and Hargreaves (2021) proved that social distancing and other safety preventative measures were met with users' positive responses.

Ⅴ. Conclusion

This study comparatively analyzed online the US tourist attractions reviews before/after the COVID 19 outbreak and mass vaccination using sentiment analysis, emotion analysis, and topic modeling techniques. We found that tourists' opinions about attractions changed from before the COVID 19 outbreak to when more than half of the local population was vaccinated. However, we did not find a significant difference in tourists' opinions before the COVID-19 outbreak and after the vaccination of half of the local population. After the COVID 19 outbreak, tourists' sentiments and happy emotions increase, and sad emotions decrease. In this way, we can see the pandemic's positive effect on tourists' opinions about attractions. Moreover, the dimension 'Safety' started to appear in tourists' opinions about attractions reviews only in the period after the outbreak and during the mass vaccination. These results mean that tourists started to care more about safety due to the impact of the COVID-19 pandemic.

The theoretical contribution of our study is that we proved that the COVID-19 outbreak and mass vaccination caused changes in tourists' opinions. Also, we introduced a methodology to evaluate tourists' opinion changes before/after the COVID-19 outbreak and the after the vaccination of more than half of the population by text mining techniques and showed the differences between period's opinions at a statistically significant level. Moreover, we proved that there were no significant differences in tourists' opinions before the pandemic and after the vaccination of half of the local population.

For tourism organizations, our study can be helpful in understanding the differences in tourists' opinions before and after the pandemic. Also, our study can improve understanding of the impact of COVID-19 on tourist behavior and may help the industry improve tourist services for higher tourist satisfaction in the new post-COVID environment.

The limitations and further research directions of our research are as follows. First, since our research aims to find only differences in the opinions of tourists from before the COVID 19 outbreak to when more than half of the local population was vaccinated, and not to explain the reason for these changes, we can only guess about the causes of these changes. Further researchers can create a study that will explain these changes in tourists' opinions caused by the COVID-19 and mass vaccination. Therefore, further research can use more reviews of different countries from other review platforms like Yelp.com, etc. Second, we calculate the sentiment scores by using the lexicon. However, one of the disadvantages of using the lexicon for sentiment analysis is that lexicon sometimes could not reflect the specific domain (Ma et al., 2018). So in further research, other approaches are needed to calculate the sentiment score of each dimension.

References

야오즈엔, 김은미, 홍태호, "온라인 리뷰의 텍스트 마이닝에 기반한 한국방문 외국인

관광객의 문화적 특성 연구,” 정보시스템연구, 제29권, 제4호, 2020, pp.171-191.

이연수, 김혜진, “소비자 리뷰 텍스트마이닝을 이용한 신생 산업 시장 구조 분석: 국내 수제 맥주 시장의 경쟁 관계 및 시장 구조를 중심으로,” 정보시스템연구, 제30권, 제2호, 2021, pp.189-214.

홍태호, 니우한잉, 임강, 박지영, “LDA를 이용한 온라인 리뷰의 다중 토픽별 감성분석- TripAdvisor 사례를 중심으로,” 정보시스템연구, 제27권, 제1호, 2018, pp.89-110.

“COVID-19: Sociocultural Impact | UNWTO.” An Inclusive Response for Vulnerable Groups, www.unwto.org/covid-19-inclusive-response-vulnerable-groups.

Aggarwal, C. C., and Zhai, C. X., “An Introduction to Text Mining,” Mining Text Data, 2012, pp. 1–10.

Alamoodi, A. H., Zaidan, B. B., Zaidan, A. A., Albahri, O. S., Mohammed, K. I., Malik, R. Q., Almahdi, E. M., Chyad, M. A., Tareq, Z., Albahri, A. S., and Hameed, H., “Sentiment Analysis and Its Applications in Fighting COVID-19 and Infectious Diseases: A Systematic Review,” Expert Systems with Applications, Vol. 167, 2020, 114155.

Al-maaitah, T. A., Majali, M. A., and Almaaitah, D. A., “The Impact of COVID-19 on the Electronic Commerce Users Behavior,” Journal of Contemporary Issues in Business and Government, Vol. 27, No.1, 2021.

Baccianella, S., Esuli, A., and Sebastiani, F., “Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining,” In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), May 2010.

Batty, M., and Taylor, M. J., “Early Processing of the Six Basic Facial Emotional Expressions,” Cognitive Brain Research, Vol. 17, No. 3, 2003, pp. 613–20.

Bernal, J. L., Andrews, N., Gower, C., Gallagher, E., Simmons, R., Thelwall, S., Stowe, J., Tessier, E., Groves, N., Dabrera, G., and Myers, R., “Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant,” New England Journal of Medicine, Vol. 385, No. 7, 2021, pp. 585–94.

Bigné, J. E., Andreu, L., and Gnoth, J., “The Theme Park Experience: An Analysis of Pleasure, Arousal and Satisfaction,” Tourism Management, Vol. 26, No. 6, 2005, pp. 833–44.

Blei, D. M., Ng, A. Y., and Jordan, M. I., “Latent Dirichlet Allocation,” Journal of Machine Learning Research, Vol. 3, January 2003, pp.993-1022.

Brown, A. M., “A New Software for Carrying out One-Way ANOVA Post Hoc

Tests," Computer Methods and Programs in Biomedicine, Vol. 79, No. 1, 2005, pp. 89–95.

Carracedo, P., Puertas, R., and Marti, L., "Research Lines on the Impact of the COVID-19 Pandemic on Business. A Text Mining Analysis," Journal of Business Research, Vol. 132, 2021, pp. 586–93.

Chua, B.-L., Al-Ansi, A., Lee, M. J., and Han, H., "Impact of Health Risk Perception on Avoidance of International Travel in the Wake of a Pandemic," Current Issues in Tourism, Vol. 24, No. 7, 2020, pp. 985–1002.

Das, D., and Bandyopadhyay, S., "Emotion Analysis on Social Media: Natural Language Processing Approaches and Applications," Online Collective Action, Springer, Vienna, 2014, pp. 19-37.

Díaz, S. S., Shaik, J. M. M., and Santofimio, J. C. G., "Intelligent Execution of Behaviors in a Nao Robot Exposed to Audiovisual Stimulus," 2018 IEEE 2nd Colombian Conference on Robotics and Automation (CCRA), November 2018, pp. 1-6.

Egger, M., Ley M., and Hanke S., "Emotion Recognition from Physiological Signal Analysis: A Review," Electronic Notes in Theoretical Computer Science, Vol. 343, 2019, pp. 35-55.

Ekman, P., "An Argument for Basic Emotions," Cognition and Emotion, Vol. 6, No. 3-4, 1992, pp. 169-200.

Fanelli, D., and Piazza, F., "Analysis and Forecast of COVID-19 Spreading in China, Italy and France," Chaos, Solitons & Fractals, Vol. 134, 2020, 109761.

Govindarajan, M., "Sentiment Analysis of Movie Reviews Using Hybrid Method of Naive Bayes and Genetic Algorithm," International Journal of Advanced Computer Research, Vol. 3, No. 4, 2013, 139.

Gretzel, U., Fuchs, M., Baggio, R., Hoepken, W., Law, R., Neidhardt, J., Pesonen, J., Zanker, M., and Xiang, Z., "E-Tourism Beyond COVID-19: a Call for Transformative Research," Information Technology & Tourism, Vol. 22, No. 2, 2020, pp.187-203.

Gursoy, D., Can, A. S., Williams, N., and Ekinci, Y., "Evolving Impacts of COVID-19 Vaccination Intentions on Travel Intentions," The Service Industries Journal, Vol. 41, No. 11–12, 2021, pp. 719-33.

Hakak, N. M., Mohd, M., Kirmani, M., and Mohd, M., "Emotion Analysis: A Survey," 2017 International Conference on Computer, Communications and Electronics (COMPTELIX), IEEE, July 2017, pp. 397-402.

Hasan, M., Agu, E., and Rundensteiner, E., "Using Hashtags as Labels for Supervised Learning of Emotions in Twitter Messages," Acm Sigkdd Workshop on Health Informatics, New York, USA, 2014.

Hu, M., and Liu, B., "Mining and Summarizing Customer Reviews," Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 2004, pp. 168-177.

Hu, N., Bose, I., Koh, N. S., and Liu, L., "Manipulation of Online Reviews: An Analysis of Ratings, Readability, and Sentiments," Decision Support Systems, Vol. 52, No. 3, 2012, pp. 674-84.

Im, J. H., Kim, J. W., and Choeh, J. Y., "COVID-19, Social Distancing, and Risk-Averse Actions of Hospitality and Tourism Consumers: A Case of South Korea," Journal of Destination Marketing & Management, Vol. 20, 2021, p. 100566.

Jacobsen, J. K. S., and Munar, A. M., "Tourist Information Search and Destination Choice in a Digital Age," Tourism Management Perspectives, Vol. 1, 2012, pp. 39-47.

Jamal, T., and Budke, C., "Tourism in a World with Pandemics: Local-Global Responsibility and Action," Journal of Tourism Futures, Vol. 6, No. 2, 2020, pp. 181-88.

Jiang, Y., and Wen, J., "Effects of COVID-19 on Hotel Marketing and Management: A Perspective Article," International Journal of Contemporary Hospitality Management, Vol. 32, No. 8, 2020, pp. 2563-73.

Khanpour, H., and Caragea, C., "Fine-grained emotion detection in health-related online posts," Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018, pp. 1160-1166.

Kumar, A., and Sebastian, T. M., "Sentiment Analysis: A Perspective on Its Past, Present and Future," International Journal of Intelligent Systems and Applications, Vol. 4, No. 10, 2012, pp. 1-14.

Li, L., Zhang, J., Nian, S., and Zhang, H., "Tourists' Perceptions of Crowding, Attractiveness, and Satisfaction: A Second-Order Structural Model," Asia Pacific Journal of Tourism Research, Vol. 22, No. 12, 2017, pp. 1250-60.

Lynott, D., and Coventry, K., "On the Ups and Downs of Emotion: Testing between Conceptual-Metaphor and Polarity Accounts of Emotional Valence–Spatial Location Interactions," Psychonomic Bulletin & Review, Vol. 21, No. 1, 2013, pp. 218-26.

Lynott, D., and Coventry, K., "On the Ups and Downs of Emotion: Testing between

Conceptual-Metaphor and Polarity Accounts of Emotional Valence-Spatial Location Interactions," Psychonomic Bulletin & Review, Vol. 21, No. 1, 2013, pp. 218-26.

Ma, E., Cheng, M., and Hsiao, A., "Sentiment Analysis - a Review and Agenda for Future Research in Hospitality Contexts," International Journal of Contemporary Hospitality Management, Vol. 30, No. 11, 2018, pp. 3287-308.

Ma, S., Sun, X., Lin, J., and Ren, X., "A Hierarchical End-to-end Model for Jointly Improving Text Summarization and Sentiment Classification," arXiv Preprint arXiv:1805.01089, 2018.

Mary, S. R., and Mahmoud, H. P., "A Model of Travel Behaviour after COVID-19 Pandemic: TripAdvisor Reviews," Current Issues in Tourism, Vol. 25, No. 7, 2022, pp. 1033-45.

Mauri, A. G., and Minazzi, R., "Web Reviews Influence on Expectations and Purchasing Intentions of Hotel Potential Customers," International Journal of Hospitality Management, Vol. 34, 2013, pp. 99-107.

Medford, R. J., Saleh, S. N., Sumarsono, A., Perl, T. M., and Lehmann, C. U., "An "Infodemic": Leveraging High-Volume Twitter Data to Understand Early Public Sentiment for the Coronavirus Disease 2019 Outbreak," Open Forum Infectious Diseases, Vol. 7, No. 7, US: Oxford University Press, 2020.

Menon, M., and Goh, T., "Tighter Border Restrictions as Coronavirus Cases in Singapore Go Past 800," The Straits Times, 2020.

Mohammad, S. M., and Turney, P. D., "Crowdsourcing a Word-Emotion Association Lexicon," Computational Intelligence, Vol. 29, No. 3, 2012, pp. 436-65.

Mukwazvure, A., and Supreethi, K. P., "A Hybrid Approach to Sentiment Analysis of News Comments," 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), IEEE, September 2015, pp. 1-6.

Munezero, M., Montero, C.S., Sutinen, E., and Pajunen, J., "Are They Different? Affect, Feeling, Emotion, Sentiment, and Opinion Detection in Text," IEEE Transactions on Affective Computing, Vol. 5, No. 2, 2014, pp. 101-111.

Ohana, B., and Tierney, B., "Sentiment Classification of Reviews Using SentiWordNet," Proceedings of IT&T, 2009.

Öztürk, N., and Ayvaz, S., "Sentiment Analysis on Twitter: A Text Mining Approach to the Syrian Refugee Crisis," Telematics and Informatics, Vol. 35, No. 1, 2018, pp. 136-47.

Pak, A., and Paroubek, P., “Twitter as a Corpus for Sentiment Analysis and Opinion Mining: Proceedings of the Seventh Conference on International Language Resources and Evaluation,” European Languages Resources Association, Valletta, Malta, 2010.

Park, I.-J., Kim, J., Kim, S., Lee, J. C., and Giroux, M., “Impact of the COVID-19 Pandemic on Travelers’ Preference for Crowded versus Non-Crowded Options,” Tourism Management, Vol. 87, 2021, p. 104398.

Pavli, A., and Helena C. M., “Travel Vaccines throughout History,” Travel Medicine and Infectious Disease, Vol. 46, 2022, p. 102278.

Poria, S., Chaturvedi, I., Cambria, E., and Bisio, F., “Sentic LDA: Improving on LDA with Semantic Similarity for Aspect-Based Sentiment Analysis,” 2016 International Joint Conference on Neural Networks (IJCNN), IEEE, July 2016, pp. 4465-4473.

Poria, S., Gelbukh, A., Hussain, A., Howard, N., Das, D., and Bandyopadhyay, S., “Enhanced SenticNet with Affective Labels for Concept-Based Opinion Mining,” IEEE Intelligent Systems, Vol. 28, No. 2, 2013, pp. 31-38.

Ram, Y., Collins-Kreiner, N., Gozansky, E., Moscona, G., and Okon-Singer, H., “Is There a COVID-19 Vaccination Effect? A Three-Wave Cross-Sectional Study,” Current Issues in Tourism, Vol. 25, No. 3, 2021, pp. 379-86.

Ren, G., and Hong, T., “Examining the Relationship between Specific Negative Emotions and the Perceived Helpfulness of Online Reviews,” Information Processing & Management, Vol. 56, No. 4, 2019, pp. 1425-38.

Ren, G., and Hong, T., “Investigating Online Destination Images Using a Topic-Based Sentiment Analysis Approach,” Sustainability, Vol. 9, No. 10, 2017, 1765.

Ridhwan, K. M., and Hargreaves, C. A., “Leveraging Twitter Data to Understand Public Sentiment for the COVID 19 Outbreak in Singapore,” International Journal of Information Management Data Insights, Vol. 1, No. 2, 2021, 100021.

Sailunaz, K., and Alhajj, R., “Emotion and Sentiment Analysis from Twitter Text,” Journal of Computational Science, Vol. 36, 2019, 101003.

Saydam, M. B., Olorunsola, V. O., Avci, T., Dambo, and T. H., Beyar, K., “How about the Service Perception during the COVID-19 Pandemic: An Analysis of Tourist Experiences from User-Generated Content on TripAdvisor,” Tourism Critiques: Practice and Theory, 2022.

Shafqat, W., and Byun, Y. C., "A Recommendation Mechanism for Under-Emphasized Tourist Spots Using Topic Modeling and Sentiment Analysis." Sustainability, Vol. 12, No. 1, 2019, p. 320.

Sharp, R. L., Sharp, J. A., and Miller, C. A., "An Island in a Sea of Development: An Examination of Place Attachment, Activity Type, and Crowding in an Urban National Park," Visitor Studies, Vol. 18, No. 2, 2015, pp. 196-213.

Shoeb, M., and Ahmed, J., "Sentiment Analysis and Classification of Tweets Using Data Mining," International Research Journal of Engineering and Technology (IRJET), Vol. 4, No. 12 2017.

Srinivasan, S. M., Sangwan, R.S., Neill, C.J., and Zu, T., "Twitter Data for Predicting Election Results: Insights from Emotion Classification," IEEE Technology and Society Magazine, Vol. 38, No. 1, 2019, pp. 58-63.

Suess, C., Maddock, J. E., Dogru, T., Mody, M., and Lee, S., "Using the Health Belief Model to Examine Travelers' Willingness to Vaccinate and Support for Vaccination Requirements Prior to Travel," Tourism Management, Vol. 88, 2022, p. 104405.

Tong, Z., and Zhang, H., "A Text Mining Research Based on LDA Topic Modelling," International Conference on Computer Science, Engineering and Information Technology, May 2016, pp. 201-210.

Tripadvisor. "Tripadvisor: Over a Billion Reviews and Contributions for Hotels, Attractions, Restaurants, and More." Tripadvisor, www.tripadvisor.com.

Williams, N. L., Nguyen, T. H. H., Del Chiappa, G., Fedeli, G., and Wassler, P., "COVID-19 Vaccine Confidence and Tourism at the Early Stage of a Voluntary Mass Vaccination Campaign: A PMT Segmentation Analysis," Current Issues in Tourism, Vol. 25, No. 3, 2021, pp. 475-89.

Xie, R., Chu, S. K. W., Chiu, D. K. W., and Wang, Y., "Exploring Public Response to COVID-19 on Weibo with LDA Topic Modeling and Sentiment Analysis," Data and Information Management, Vol. 5, No. 1, 2021, pp. 86-99.

Yang, H., Willis, A., De Roeck, A., and Nuseibeh, B., "A Hybrid Model for Automatic Emotion Recognition in Suicide Notes," Biomedical Informatics Insights, Vol. 5s1, 2012, p. BII.S8948.

Yoon, J. I., Kyle, G., Hsu, Y.-C., and Absher, J., "Coping with Crowded Recreation Settings: A Cross-Cultural Investigation," Journal of Leisure Research, Vol. 52, No. 1, 2020, pp. 1-21.

Zaman, U., Aktan, M., Anjam, M., Agrusa, J., Khwaja, M. G., and Farías, P., "Can Post-Vaccine 'Vaxication' Rejuvenate Global Tourism? Nexus between COVID-19 Branded Destination Safety, Travel Shaming, Incentives and the Rise of Vaxication Travel," Sustainability, Vol. 13, No. 24, 2021, p. 14043.

Zhao, W., Chen, J. J., Perkins, R., Liu, Z., Ge, W., Ding, Y., and Zou, W., "A Heuristic Approach to Determine an Appropriate Number of Topics in Topic Modeling." BMC Bioinformatics, Vol. 16, No. 13, 2015, pp. 1-10.

Chernyaeva, Olga

Olga, Chernyaeva is a Ph.D. student of Management Information Systems at College of Business Administration, Pusan National University in Korea. She received a Master's degree from Pusan National University. Her research interest includes business analytics, intelligent systems, data mining, and recommender systems for e-business.

Ziyan, Yao

Ziyan, Yao is a Ph.D. student of Management Information Systems at College of Business Administration, Pusan National University in Korea. He received a Master's degree from Pusan National University. His research interest includes eWOM, opinion mining, and social network.

Hong, Tae Ho

Taeho, Hong is a Professor of Management Information Systems at College of Business Administration, Pusan National University in Korea. He received a Ph.D. from KAIST. His research interest includes intelligent systems, data mining, and recommender systems for e-business. He has published his research in Expert Systems with Application, Expert Systems, Information Processing & Management, and many other journals.

<Abstract>

Understanding the Changes in Tourists' Opinions in the Era of the COVID-19

Chernyaeva, Olga · Ziyan, Yao · Hong, Tae Ho

Purpose

The purpose of this study is to explore and compare changes in tourist opinion during the COVID-19 pandemic. Since the COVID-19 outbreak has caused changes in all areas of our lives, the conditions related to confinement during a lockdown have led to changes in tourists' habits and behaviors.

Design/methodology/approach

To analyze opinion changes about tourist attractions, this study performed topic modeling by summarizing topics into five dimensions: management, scenery, price, suggestion, and safety; then, based on the topic modeling results, sentiment analysis and emotion analysis were conducted to explore the change of tourists' opinion during the COVID-19 pandemic.

Findings

According to the results, this study confirmed the pandemic's positive effect on tourists' opinions about attractions after the COVID 19 outbreak. Presumably due to the absence of lines and crowed. Moreover, the dimension 'Safety' started to appear in US tourists' attractions reviews only in the period after the outbreak and during the mass vaccination. These results mean that tourists started to care more about safety due to the impact of the COVID-19 pandemic.

Keyword: COVID-19, Online Customers Reviews, Topic Modeling, Sentiment Analysis, Emotion Analysis, Tourist Attractions

* 이 논문은 2022년 4월 12일 접수, 2022년 5월 19일 1차 심사, 2022년 6월 24일 게재 확정되었습니다.